Week 52

Shaun Stearns

9/26/2019

Multiple Regression 1

Lowest Weigh-in

```
DF <- read.csv("C:/Users/LaoTz/Desktop/DF Articles/WeightLoss.csv", header =
TRUE)
DF <- na.omit(DF)
DF.t <- DF[-c(36,37,56)]
DF.t <- DF.t[c(10,20,23,26,29,32,35,38,41,44,47,50,53)]
DF.t = scale(DF.t, center = TRUE, scale = TRUE)
DF.t <- as.data.frame(DF.t)
reg <- lm(Lowestweighinkg ~., DF.t)
```

```
MODEL INFO:
Observations: 7138
Dependent Variable: Lowestweighinkg
Type: OLS linear regression

MODEL FIT:
F(12,7125) = 26.49, p = 0.00
R² = 0.04
Adj. R² = 0.04

Standard errors: OLS
```

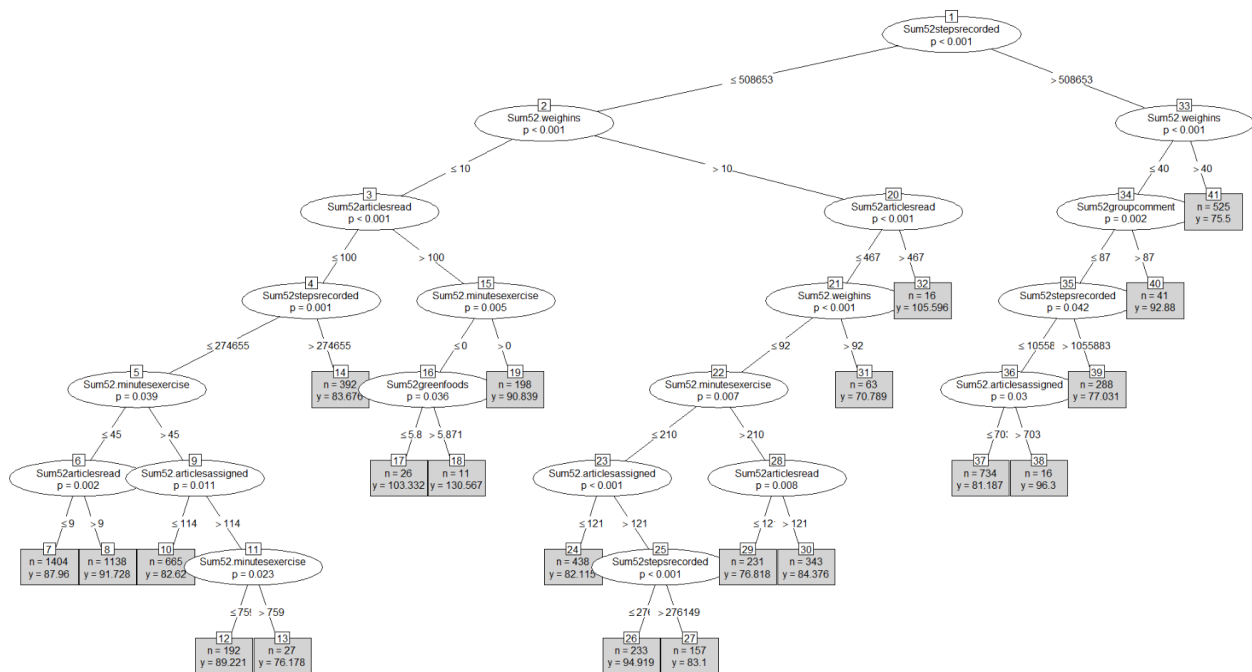|	Est.	2.5%	97.5%	t val.	p	partial.r	part.r
(Intercept)	0.00	-0.02	0.02	0.00	1.00		
Sum52mealslogged	-0.10	-0.16	-0.04	-3.17	0.00	-0.04	-0.04
Sum52greenfoods	0.06	0.02	0.09	3.34	0.00	0.04	0.04
Sum52exerciselogged	-0.00	-0.04	0.04	-0.11	0.91	-0.00	-0.00
Sum52.minutesexercise	-0.06	-0.10	-0.02	-3.23	0.00	-0.04	-0.04
Sum52stepsrecorded	-0.16	-0.19	-0.13	-9.73	0.00	-0.11	-0.11
Sum52.weighins	-0.11	-0.14	-0.09	-8.00	0.00	-0.09	-0.09
Sum52.articlesassigned	0.07	0.01	0.12	2.16	0.03	0.03	0.03
Sum52articlesread	0.10	0.05	0.16	3.88	0.00	0.05	0.05
Sum52grouppost	0.03	-0.00	0.06	1.83	0.07	0.02	0.02
Sum52groupcomment	0.01	-0.03	0.04	0.33	0.74	0.00	0.00
Sum52grouplikes	-0.00	-0.04	0.03	-0.08	0.94	-0.00	-0.00
Sum52messagestocoach	0.01	-0.02	0.04	0.48	0.63	0.01	0.01
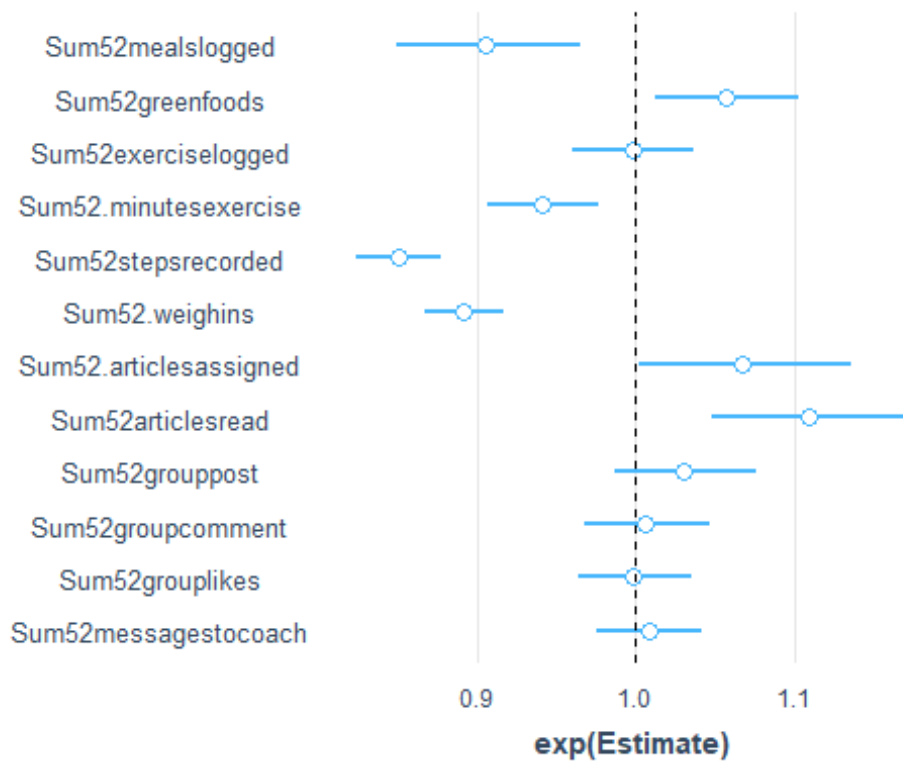
#Regression Tree 1

```r
DF <- read.csv("C:/Users/LaoTz/Desktop/DF Articles/WeightLoss.csv", header =
TRUE)
DF <- na.omit(DF)
DF.t <- DF[-c(36,37,56)]
DF.t <- DF.t[c(10,20,23,26,29,32,35,38,41,44,47,50,53)]
model <- train(
  Lowestweighinkg ~., DF.t, method = "ctree",
  trControl = trainControl("cv", number = 10),
  tuneGrid = expand.grid(mincriterion = 0.95)
)
model$results

##   mincriterion     RMSE  Rsquared      MAE    RMSESD RsquaredSD      MAESD
## 1         0.95 20.19509 0.0396779 15.61248 0.6364691 0.01547453 0.3528991
```

Tree Model

```r
plot(model$finalModel, type = "simple")
```

Sum52stepsrecorded
p < 0.001

≤ 508653 | > 508653

Sum52 weighins
p < 0.001

Sum52 weighins
p < 0.001

≤ 10 | > 10

≤ 40 | > 40

Sum52articlesread
p < 0.001

Sum52articlesread
p < 0.001

Sum52groupcomment
p = 0.002

n = 525
y = 75.5

≤ 100 | > 100

≤ 467 | > 467

≤ 87 | > 87

Sum52stepsrecorded
p = 0.001

Sum52.minutesexercise
p = 0.005

Sum52 weighins
p < 0.001

n = 16
y = 105.596

Sum52stepsrecorded
p = 0.042

n = 41
y = 92.88

≤ 274655 | > 274655

≤ 0 | > 0

≤ 92 | > 92

≤ 1055885 | > 1055883

Sum52.minutesexercise
p = 0.039

n = 392
y = 83.676

Sum52greenfoods
p = 0.036

n = 198
y = 90.839

Sum52.minutesexercise
p = 0.007

n = 63
y = 70.789

Sum52 articlesassigned
p = 0.03

n = 288
y = 77.031

≤ 45 | > 45

≤ 5.8 | > 5.871

≤ 210 | > 210

≤ 703 | > 703

Sum52articlesread
p = 0.002

Sum52 articlesassigned
p = 0.011

n = 26
y = 103.332

n = 11
y = 130.567

Sum52 articlesassigned
p < 0.001

Sum52articlesread
p = 0.008

n = 734
y = 81.187

n = 16
y = 96.3

≤ 9 | > 9

≤ 114 | > 114

≤ 121 | > 121

≤ 12 | > 121

n = 1404
y = 87.96

n = 1138
y = 91.728

n = 665
y = 82.62

Sum52.minutesexercise
p = 0.023

n = 438
y = 82.115

Sum52stepsrecorded
p < 0.001

n = 231
y = 76.818

n = 343
y = 84.376

≤ 759 | > 759

≤ 276 | > 276149

n = 192
y = 89.221

n = 27
y = 76.178

n = 233
y = 94.919

n = 157
y = 83.1

User Engagement and Lowest Weigh-in based on 22 Terminal Nodes (Left to Right)

1	Lowest Weigh-in avg 87.96 kg **= WK 52 Steps Recorded ≤ 508653, WK 52 Weigh-ins ≤ 10, WK 52 Articles Read ≤ 100, WK 52 Steps Recorded ≤ 274655, WK 52 Min of Exercise ≤ 45, WK 52 Articles Read ≤ 9**
2	Lowest Weigh-in avg 91.73 kg = **WK 52 Steps Recorded ≤ 508653, WK 52 Weigh-ins ≤ 10, WK 52 Articles Read ≤ 100, WK 52 Steps Recorded ≤ 274655, WK 52 Min of Exercise ≤ 45, WK 52 Articles Read > 9**
3	Lowest Weigh-in avg 82.62 kg = **WK 52 Steps Recorded ≤ 508653, WK 52 Weigh-ins ≤ 10, WK 52 Articles Read ≤ 100, WK 52 Steps Recorded ≤ 274655, WK 52 Min of Exercise > 45, WK 52 Articles Assigned ≤ 114**
4	Lowest Weigh-in avg 89.22 kg = **WK 52 Steps Recorded ≤ 508653, WK 52 Weigh-ins ≤ 10, WK 52 Articles Read ≤ 100, WK 52 Steps Recorded ≤ 274655, WK 52 Min of Exercise > 45, WK 52 Articles Assigned > 114, WK 52 Min of Exercise ≤ 759**
5	Lowest Weigh-in avg 76.18 kg = **WK 52 Steps Recorded ≤ 508653, WK 52 Weigh-ins ≤ 10, WK 52 Articles Read ≤ 100, WK 52 Steps Recorded ≤ 274655, WK 52 Min of Exercise > 45, WK 52 Articles Assigned > 114, WK 52 Min of Exercise >759**
6	Lowest Weigh-in avg 83.68 kg = **WK 52 Steps Recorded ≤ 508653, WK 52 Weigh-ins ≤ 10, WK 52 Articles Read ≤ 100, WK 52 Steps Recorded > 274655**
7	Lowest Weigh-in avg 103.33 kg = **WK 52 Steps Recorded ≤ 508653, WK 52 Weigh-ins ≤ 10, WK 52 Articles Read > 100, WK 52 Min of Exercise ≤ 0, WK 52 Green Foods ≤ 5.8**
8 2nd High	Lowest Weigh-in avg 103.57 kg = **WK 52 Steps Recorded ≤ 508653, WK 52 Weigh-ins ≤ 10, WK 52 Articles Read > 100, WK 52 Min of Exercise ≤ 0, WK 52 Green Foods > 5.8**
9	Lowest Weigh-in avg 90.84 kg = **WK 52 Steps Recorded ≤ 508653, WK 52 Weigh-ins ≤ 10, WK 52 Articles Read > 100, WK 52 Min of Exercise > 0**
10	Lowest Weigh-in avg 82.12 kg = **WK 52 Steps Recorded ≤ 508653, WK 52 Weigh-ins > 10, WK 52 Articles Read ≤ 467, WK 52 Weigh-ins ≤ 92, WK 52 Min of Exercise ≤ 210, WK 52 Articles Assigned ≤ 121**

11	Lowest Weigh-in avg 94.92 kg = **WK 52 Steps Recorded $\leq$ 508653, WK 52 Weigh-ins > 10, WK 52 Articles Read $\leq$ 467, WK 52 Weigh-ins $\leq$ 92, WK 52 Min of Exercise $\leq$ 210, WK 52 Articles Assigned > 121, WK 52 Steps Recorded $\leq$ 276149**
12	Lowest Weigh-in avg 83.1 kg = **WK 52 Steps Recorded $\leq$ 508653, WK 52 Weigh-ins > 10, WK 52 Articles Read $\leq$ 467, WK 52 Weigh-ins $\leq$ 92, WK 52 Min of Exercise $\leq$ 210, WK 52 Articles Assigned > 121, WK 52 Steps Recorded > 276149**
13	Lowest Weigh-in avg 76.82 kg = **WK 52 Steps Recorded $\leq$ 508653, WK 52 Weigh-ins > 10, WK 52 Articles Read $\leq$ 467, WK 52 Weigh-ins $\leq$ 92, WK 52 Min of Exercise $\leq$ 210, WK 52 Articles Read $\leq$ 121**
14	Lowest Weigh-in avg 84.38 kg = **WK 52 Steps Recorded $\leq$ 508653, WK 52 Weigh-ins > 10, WK 52 Articles Read $\leq$ 467, WK 52 Weigh-ins $\leq$ 92, WK 52 Min of Exercise $\leq$ 210, WK 52 Articles Read > 121**
15 Low	Lowest Weigh-in avg 70.79 kg = **WK 52 Steps Recorded $\leq$ 508653, WK 52 Weigh-ins > 10, WK 52 Articles Read $\leq$ 467, WK 52 Weigh-ins > 92**
16 High	Lowest Weigh-in avg 105.57 kg = **WK 52 Steps Recorded $\leq$ 508653, WK 52 Weigh-ins > 10, WK 52 Articles Read > 467**
17	Lowest Weigh-in avg 81.19 kg = **WK 52 Steps Recorded > 508653, WK 52 Weigh-ins $\leq$ 40, WK 52 Group Comments $\leq$ 87, WK 52 Steps Recorded $\leq$ 1055883, WK 52 Articles Assigned $\leq$ 703**
18	Lowest Weigh-in avg 96.3 kg = **WK 52 Steps Recorded > 508653, WK 52 Weigh-ins $\leq$ 40, WK 52 Group Comments $\leq$ 87, WK 52 Steps Recorded $\leq$ 1055883, WK 52 Articles Assigned > 703**
19	Lowest Weigh-in avg 77.03 kg = **WK 52 Steps Recorded > 508653, WK 52 Weigh-ins $\leq$ 40, WK 52 Group Comments $\leq$ 87, WK 52 Steps Recorded > 1055883**
20	Lowest Weigh-in avg 92.88 kg = **WK 52 Steps Recorded > 508653, WK 52 Weigh-ins $\leq$ 40, WK 52 Group Comments > 87**
21	Lowest Weigh-in avg 75.5 kg = **WK 52 Steps Recorded > 508653, WK 52 Weigh-ins > 40**

Random Forest 1

```
DF <- read.csv("C:/Users/LaoTz/Desktop/DF Articles/WeightLoss.csv", header =
TRUE)
DF <- na.omit(DF)
DF.t <- DF[-c(36,37,56)]
DF.t <- DF.t[c(10,20,23,26,29,32,35,38,41,44,47,50,53)]
rf <- randomForest(Lowestweighinkg ~ ., data = DF.t, ntree = 25,
                   mtry = 4, nodesize = 5, importance = TRUE)
```
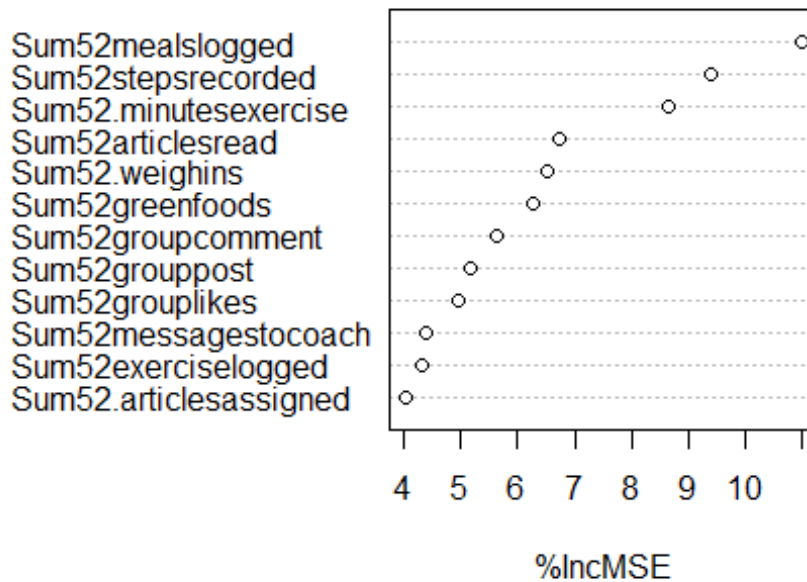
Signicance Testing

```
rf.perm <- rf.significance(rf, DF.t, q = 0.99, p = 0.05, nperm=99, ntree=25)
rf.perm

## Number of permutations:  99
## p-value:  0.01
## Model signifiant at p = 0.01
##    Model R-square:  -0.1254238
##    Random R-square:  -0.2659611
##    Random R-square variance:  0.0001811125
```

Variable Importance Plot

```
varImpPlot(rf, type = 1, main = "Lowest Rec Weight")
```

Lowest Rec Weight

%IncMSE

Multipleression 2

Difference Between First Weigh-in and Lowest Weigh-in

```
DF <- read.csv("C:/Users/LaoTz/Desktop/DF Articles/WeightLoss.csv", header =
TRUE)
DF <- na.omit(DF)
DF.t <- DF[-c(36,37,56)]
DF.t <- DF.t[c(11,20,23,26,29,32,35,38,41,44,47,50,53)]
DF.t = scale(DF.t, center = TRUE, scale = TRUE)
DF.t <- as.data.frame(DF.t)
reg <- lm(AbsDiffFirstWeighinkg ~., DF.t)
```

```
MODEL INFO:
Observations: 7138
Dependent Variable: AbsDiffFirstWeighinkg
Type: OLS linear regression

MODEL FIT:
F(12,7125) = 71.86, p = 0.00
R² = 0.11
Adj. R² = 0.11

Standard errors: OLS
```

	Est.	2.5%	97.5%	t val.	p	partial.r	part.r
(Intercept)	-0.00	-0.02	0.02	-0.00	1.00		
Sum52mealslogged	-0.01	-0.07	0.05	-0.22	0.82	-0.00	-0.00
Sum52greenfoods	-0.01	-0.04	0.02	-0.47	0.64	-0.01	-0.01
Sum52exerciselogged	-0.03	-0.07	0.01	-1.50	0.13	-0.02	-0.02
Sum52.minutesexercise	-0.01	-0.04	0.03	-0.40	0.69	-0.00	-0.00
Sum52stepsrecorded	0.04	0.01	0.07	2.51	0.01	0.03	0.03
Sum52.weighins	0.26	0.23	0.29	18.61	0.00	0.22	0.21
Sum52.articlesassigned	0.16	0.10	0.21	5.38	0.00	0.06	0.06
Sum52articlesread	-0.04	-0.09	0.01	-1.61	0.11	-0.02	-0.02
Sum52grouppost	-0.00	-0.03	0.03	-0.14	0.89	-0.00	-0.00
Sum52groupcomment	0.02	-0.02	0.06	1.14	0.25	0.01	0.01
Sum52grouplikes	0.02	-0.02	0.05	1.03	0.30	0.01	0.01
Sum52messagestocoach	-0.05	-0.08	-0.02	-3.28	0.00	-0.04	-0.04
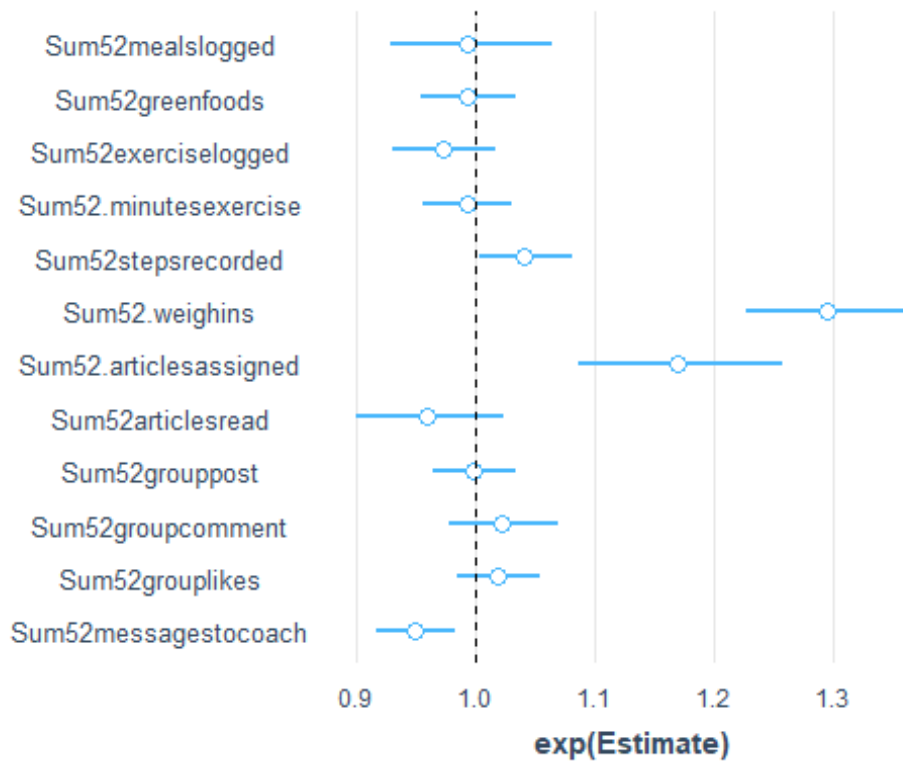
#Regression Tree 2

```
DF <- read.csv("C:/Users/LaoTz/Desktop/DF Articles/WeightLoss.csv", header =
TRUE)
DF <- na.omit(DF)
```

```
DF.t <- DF[-c(36,37,56)]
DF.t <- DF.t[c(11,20,23,26,29,32,35,38,41,44,47,50,53)]
model <- train(
  AbsDiffFirstWeighinkg ~., DF.t, method = "ctree",
  trControl = trainControl("cv", number = 10),
  tuneGrid = expand.grid(mincriterion = 0.95)
)
model$results

##   mincriterion     RMSE  Rsquared      MAE   RMSESD RsquaredSD      MAESD
## 1         0.95 7.607899 0.1156952 3.793135 1.112919 0.02329779 0.2332627
```

Tree Plot

```
plot(model$finalModel, type = "simple")
```
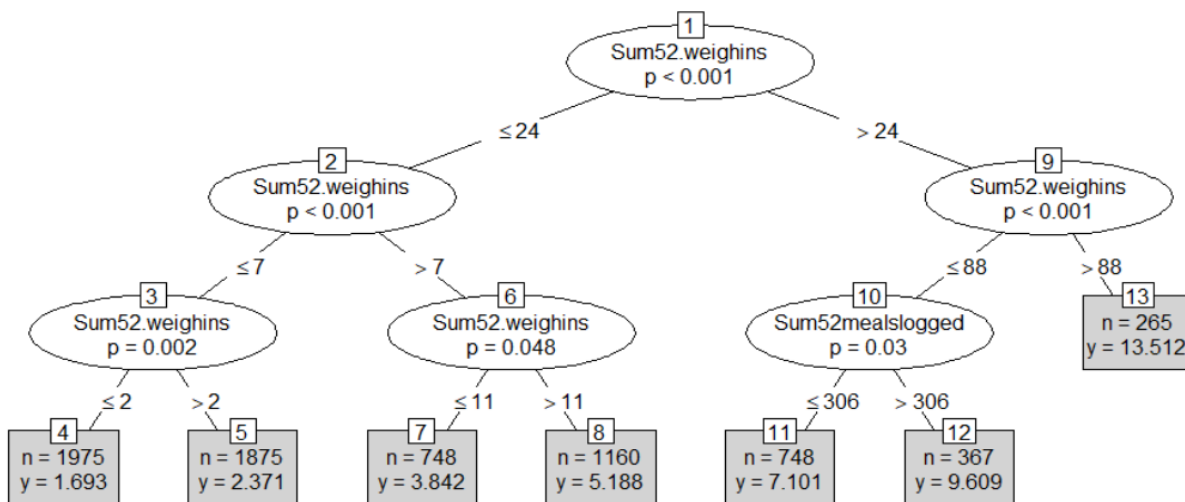

User Engagement and Difference Between First Weigh-in and Lowest Weigh-in based on 7 Terminal Nodes (Left to Right)

1 High	Diff 1st Weight Lost avg 1.69 kg = **WK 52 Weigh-ins ≤ 2**
2	Diff 1st Weight Lost avg 2.37 kg = **WK 52 Weigh-ins > 2 & ≤ 7**
3	Diff 1st Weight Lost avg 3.84 kg = **WK 52 Weigh-ins > 7 & ≤ 11**
4	Diff 1st Weight Lost avg 5.19 kg = **WK 52 Weigh-ins > 7 & > 11**
5	Diff 1st Weight Lost avg 7.10 kg = **WK 52 Weigh-ins > 24, WK 52 Weigh-ins ≤ 88, WK 52 Meals Logged ≤ 306**
6	Diff 1st Weight Lost avg 9.61 kg = **WK 52 Weigh-ins > 52, WK 52 Weigh-ins ≤ 88, WK 52 Meals Logged > 306**
7 Low	Diff 1st Weight Lost avg 13.51 kg = **WK 52 Weigh-ins > 88**

Random Forest 2

```
DF <- read.csv("C:/Users/LaoTz/Desktop/DF Articles/WeightLoss.csv", header =
TRUE)
DF <- na.omit(DF)
DF.t <- DF[-c(36,37,56)]
DF.t <- DF.t[c(11,20,23,26,29,32,35,38,41,44,47,50,53)]
rf <- randomForest(AbsDiffFirstWeighinkg ~ ., data = DF.t, ntree = 25,
                  mtry = 4, nodesize = 5, importance = TRUE)
```
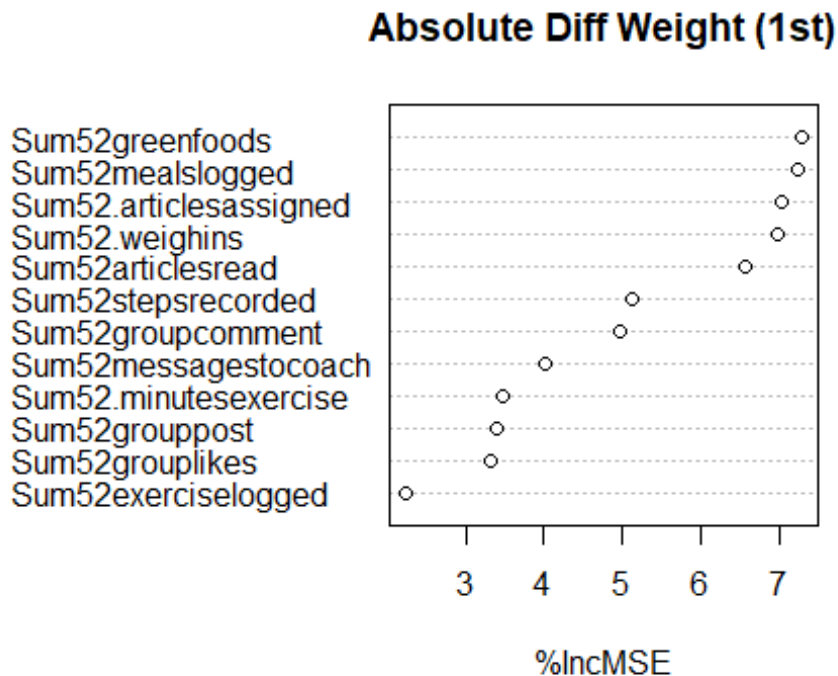
Signicance Testing

```
rf.perm <- rf.significance(rf, DF.t, q = 0.99, p = 0.05, nperm=99, ntree=25)
rf.perm

## Number of permutations:  99
## p-value:  0.01
## Model signifiant at p = 0.01
##    Model R-square:  -0.09286596
##    Random R-square:  -0.2799836
##    Random R-square variance:  0.0004081632
```

Variable Importance Plot

```
varImpPlot(rf, type = 1, main = "Absolute Diff Weight (1st)")
```

Multipleression 3

Differnece Between Initial Weigh-in and Lowest Weigh-in

```
DF <- read.csv("C:/Users/LaoTz/Desktop/DF Articles/WeightLoss.csv", header =
TRUE)
DF <- na.omit(DF)
DF.t <- DF[-c(36,37,56)]
DF.t <- DF.t[c(12,20,23,26,29,32,35,38,41,44,47,50,53)]
DF.t = scale(DF.t, center = TRUE, scale = TRUE)
DF.t <- as.data.frame(DF.t)
reg <- lm(AbsDiffInitWeighinkg ~., DF.t)
```

MODEL INFO:
Observations: 7138
Dependent Variable: AbsDiffInitWeighinkg
Type: OLS linear regression

MODEL FIT:
$F_{(12,7125)}$ = 60.59, p = 0.00
R^2 = 0.09
Adj. R^2 = 0.09

Standard errors: OLS

|	Est.	2.5%	97.5%	t val.	p	partial.r	part.r
(Intercept)	-0.00	-0.02	0.02	-0.00	1.00		
Sum52mealslogged	-0.05	-0.11	0.01	-1.56	0.12	-0.02	-0.02
Sum52greenfoods	-0.00	-0.03	0.03	-0.19	0.85	-0.00	-0.00
Sum52exerciselogged	-0.05	-0.09	-0.01	-2.69	0.01	-0.03	-0.03
Sum52.minutesexercise	0.00	-0.03	0.04	0.20	0.84	0.00	0.00
Sum52stepsrecorded	0.05	0.02	0.09	3.32	0.00	0.04	0.04
Sum52.weighins	0.25	0.22	0.28	17.95	0.00	0.21	0.20
Sum52.articlesassigned	0.14	0.08	0.20	4.78	0.00	0.06	0.05
Sum52articlesread	-0.01	-0.06	0.04	-0.51	0.61	-0.01	-0.01
Sum52grouppost	-0.01	-0.05	0.02	-0.91	0.36	-0.01	-0.01
Sum52groupcomment	0.02	-0.02	0.05	0.91	0.36	0.01	0.01
Sum52grouplikes	0.02	-0.02	0.05	0.87	0.38	0.01	0.01
Sum52messagestocoach	-0.04	-0.07	-0.01	-2.37	0.02	-0.03	-0.03
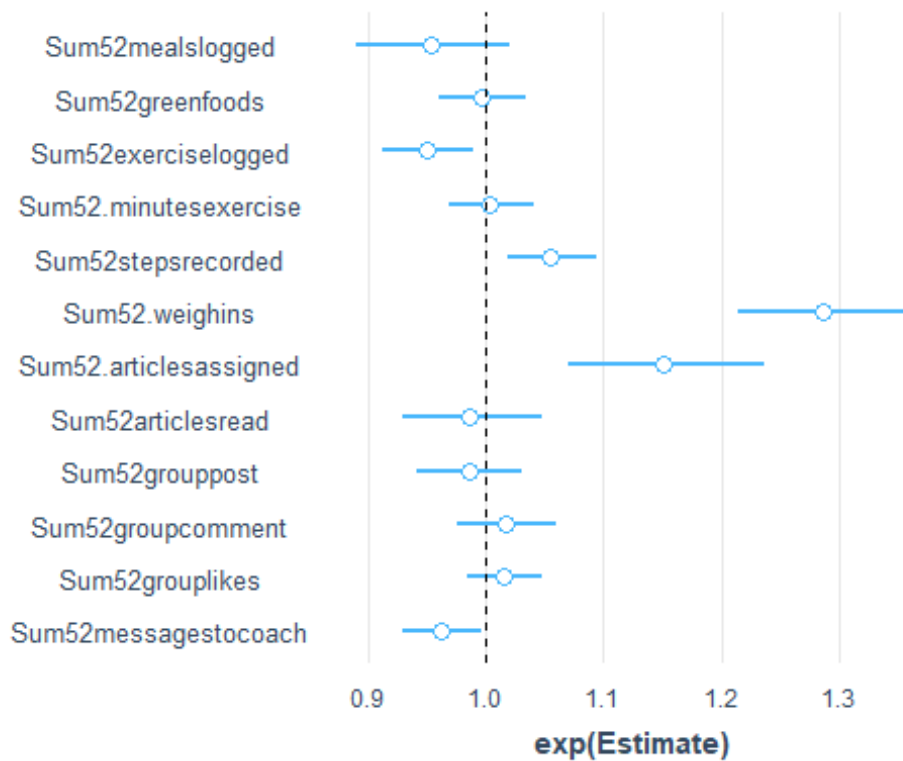
#Regression Tree 3

```
DF <- read.csv("C:/Users/LaoTz/Desktop/DF Articles/WeightLoss.csv", header =
TRUE)
DF <- na.omit(DF)
DF.t <- DF[-c(36,37,56)]
DF.t <- DF.t[c(12,20,23,26,29,32,35,38,41,44,47,50,53)]
model <- train(
  AbsDiffInitWeighinkg ~., DF.t, method = "ctree",
  trControl = trainControl("cv", number = 10),
  tuneGrid = expand.grid(mincriterion = 0.95)
)
model$results

##   mincriterion     RMSE  Rsquared      MAE    RMSESD RsquaredSD      MAESD
## 1         0.95 8.315914 0.1082704 4.274393 0.7375855 0.04088306 0.1878762
```

Tree Model

```
plot(model$finalModel, type = "simple")
```
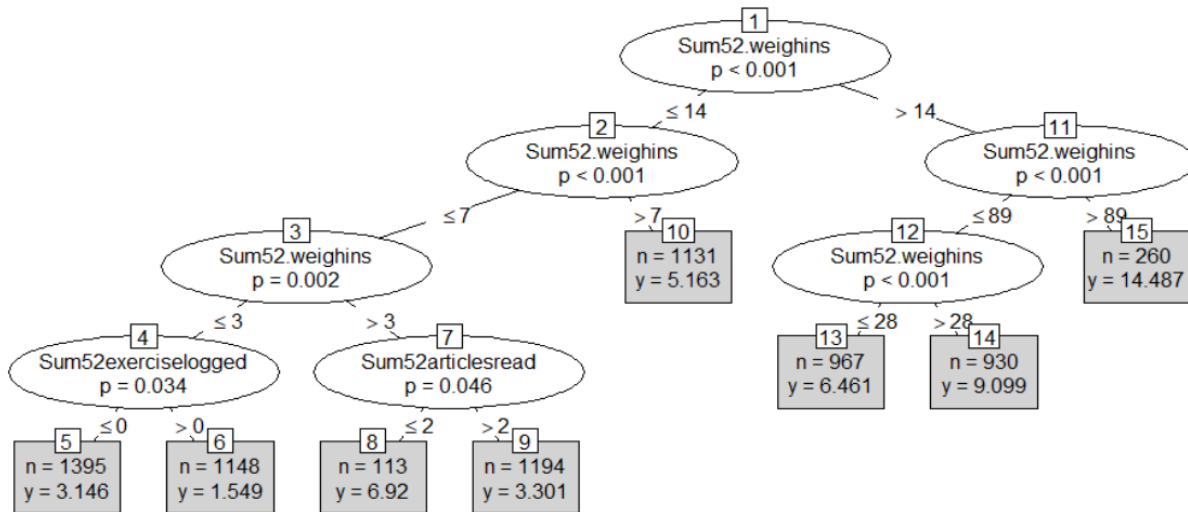
User Engagement and Difference Between Initial Weigh-in and Lowest Weigh-in based on 8 Terminal Nodes (Left to Right)

1	Diff Initial Weight Lost avg 3.15 kg = **WK 52 Weigh-ins ≤ 3, WK 52 Exercise Logged ≤ 0**
2 High	Diff Initial Weight Lost avg 1.55 kg = **WK 52 Weigh-ins ≤ 3, WK 52 Exercise Logged > 0**
3	Diff Initial Weight Lost avg 6.92 kg = **WK 52 Weigh-ins ≤ 7, WK 52 Weigh-ins > 3, WK 52 Articles Read ≤ 2**
4	Diff Initial Weight Lost avg 3.30 kg = **WK 52 Weigh-ins ≤ 7, WK 52 Weigh-ins > 3, WK 52 Articles Read > 2**
5	Diff Initial Weight Lost avg 5.16 kg = **WK52 Weigh-ins ≤ 14, WK 52 Weigh-ins > 7**
6	Diff Initial Weight Lost avg 6.46 kg = **WK 52 Weigh-ins > 14, WK 52 Weigh-ins ≤ 28**
7	Diff Initial Weight Lost avg 9.10 kg = **WK 24 Weigh-ins > 27, WK 24 Weigh-ins ≤ 89, WK 52 Weigh-ins > 28**
8 Low	Diff Initial Weight Lost avg 14.49 kg = **WK 24 Weigh-ins > 27, WK 24 Weigh-ins > 89**

Random Forest 3

```
DF <- read.csv("C:/Users/LaoTz/Desktop/DF Articles/WeightLoss.csv", header =
TRUE)
DF <- na.omit(DF)
DF.t <- DF[-c(36,37,56)]
DF.t <- DF.t[c(12,20,23,26,29,32,35,38,41,44,47,50,53)]
rf <- randomForest(AbsDiffInitWeighinkg ~ ., data = DF.t, ntree = 25,
                   mtry = 4, nodesize = 5, importance = TRUE)
```
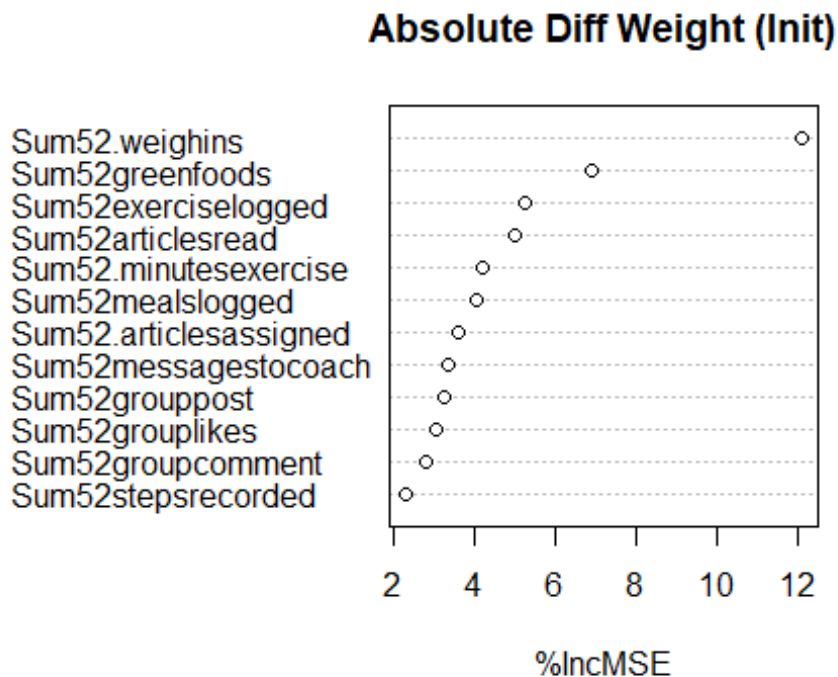
Signicance Testing

```
rf.perm <- rf.significance(rf, DF.t, q = 0.99, p = 0.05, nperm=99, ntree=25)
rf.perm

## Number of permutations:  99
## p-value:  0.01
```

```
## Model signifiant at p = 0.01
##   Model R-square:  -0.0825142
##   Random R-square:  -0.2744248
##   Random R-square variance:  0.0005864191
```

Variable Importance Plot

```
varImpPlot(rf, type = 1, main = "Absolute Diff Weight (Init)")
```


Absolute Diff Weight (Init)

Multipleression 4

Curriculum Week (or Length of Time with DF?)

```
DF <- read.csv("C:/Users/LaoTz/Desktop/DF Articles/WeightLoss.csv", header =
TRUE)
DF <- na.omit(DF)
DF.t <- DF[-c(36,37,56)]
DF.t <- DF.t[c(17,20,23,26,29,32,35,38,41,44,47,50,53)]
DF.t = scale(DF.t, center = TRUE, scale = TRUE)
DF.t <- as.data.frame(DF.t)
reg <- lm(CurriculumWeek ~., DF.t)
```

MODEL INFO:
Observations: 7138
Dependent Variable: CurriculumWeek
Type: OLS linear regression

MODEL FIT:
$F(12,7125) = 6.68$, $p = 0.00$
$R^2 = 0.01$
Adj. $R^2 = 0.01$

Standard errors: OLS

|	Est.	2.5%	97.5%	t val.	p	partial.r	part.r
(Intercept)	-0.00	-0.02	0.02	-0.00	1.00		
Sum52mealslogged	0.11	0.05	0.17	3.47	0.00	0.04	0.04
Sum52greenfoods	0.00	-0.03	0.04	0.17	0.86	0.00	0.00
Sum52exerciselogged	0.03	-0.01	0.07	1.51	0.13	0.02	0.02
Sum52.minutesexercise	-0.04	-0.07	0.00	-1.85	0.06	-0.02	-0.02
Sum52stepsrecorded	0.00	-0.03	0.03	0.09	0.93	0.00	0.00
Sum52.weighins	-0.02	-0.05	0.00	-1.71	0.09	-0.02	-0.02
Sum52.articlesassigned	-0.04	-0.10	0.02	-1.44	0.15	-0.02	-0.02
Sum52articlesread	-0.10	-0.15	-0.05	-3.66	0.00	-0.04	-0.04
Sum52grouppost	0.06	0.02	0.09	3.27	0.00	0.04	0.04
Sum52groupcomment	-0.02	-0.06	0.02	-0.97	0.33	-0.01	-0.01
Sum52grouplikes	0.01	-0.03	0.05	0.56	0.57	0.01	0.01
Sum52messagestocoach	0.09	0.06	0.12	5.42	0.00	0.06	0.06
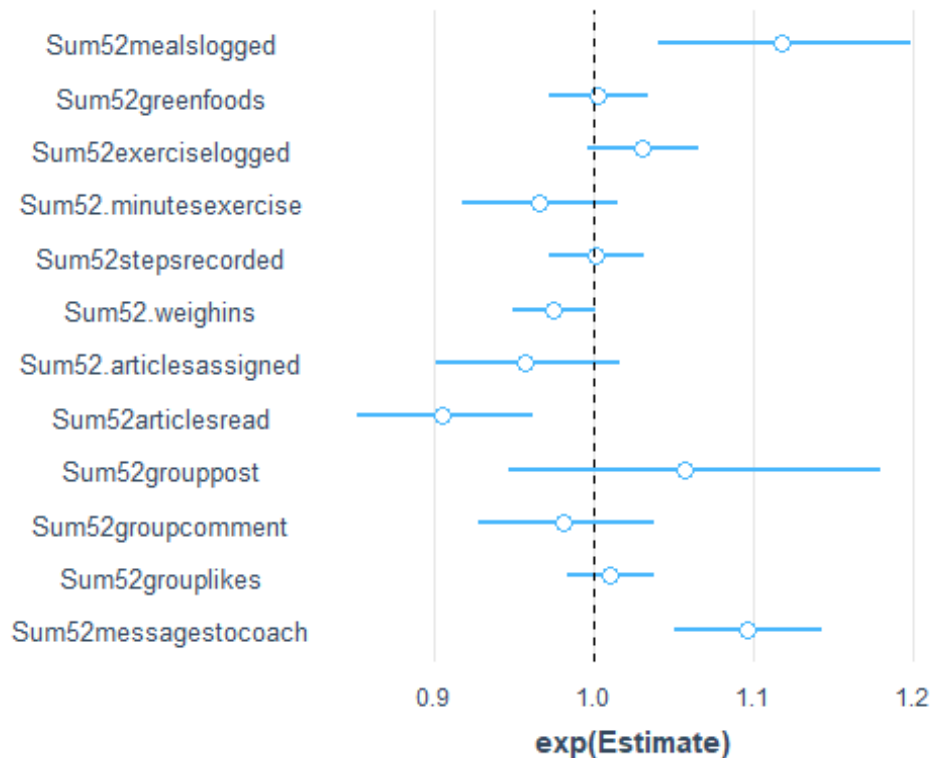
#Regression Tree 4

```r
DF <- read.csv("C:/Users/LaoTz/Desktop/DF Articles/WeightLoss.csv", header =
TRUE)
DF <- na.omit(DF)
DF.t <- DF[-c(36,37,56)]
DF.t <- DF.t[c(17,20,23,26,29,32,35,38,41,44,47,50,53)]
model <- train(
  CurriculumWeek ~., DF.t, method = "ctree",
  trControl = trainControl("cv", number = 10),
  tuneGrid = expand.grid(mincriterion = 0.95)
)
model$results
```

```
##   mincriterion     RMSE   Rsquared      MAE   RMSESD RsquaredSD     MAESD
## 1         0.95 14.69248 0.01000034 8.344469 1.043133 0.01641986 0.4284138
```

Tree Model

```r
plot(model$finalModel, type = "simple")
```
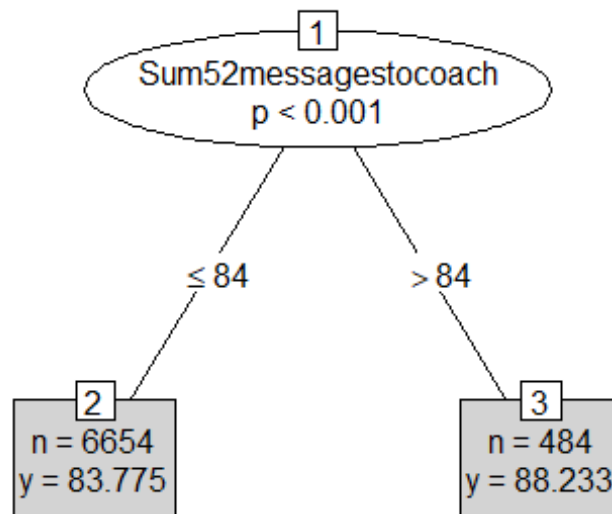

User Engagement and Curriculum Week based on 2 Terminal Nodes (Left to Right)

1	Curriculum Week avg 83.78 = **WK 52 Messages to Coach $\leq$ 84**
2	Curriculum Week avg 88.23 = **WK 52 Messages to Coach > 84**

Random Forest 4

```r
DF <- read.csv("C:/Users/LaoTz/Desktop/DF Articles/WeightLoss.csv", header =
TRUE)
DF <- na.omit(DF)
DF.t <- DF[-c(36,37,56)]
DF.t <- DF.t[c(17,20,23,26,29,32,35,38,41,44,47,50,53)]
rf <- randomForest(CurriculumWeek ~ ., data = DF.t, ntree = 25,
                   mtry = 4, nodesize = 5, importance = TRUE)
```
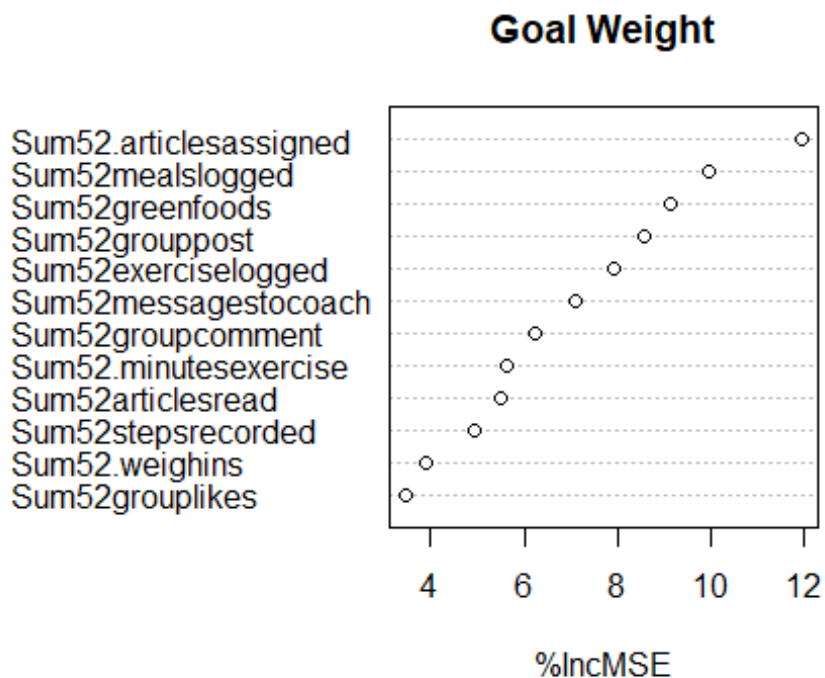
Signicance Testing

```r
rf.perm <- rf.significance(rf, DF.t, q = 0.99, p = 0.05, nperm=99, ntree=25)
rf.perm

## Number of permutations:  99
## p-value:  0.01
## Model signifiant at p = 0.01
##    Model R-square:  -0.08634552
##    Random R-square:  -0.2752181
##    Random R-square variance:  0.0003134305
```

Variable Importance Plot

```r
varImpPlot(rf, type = 1, main = "Goal Weight")
```

Multipleression 5

Differnce in BMI

```r
DF <- read.csv("C:/Users/LaoTz/Desktop/DF Articles/WeightLoss.csv", header =
TRUE)
DF <- na.omit(DF)
DF.t <- DF[-c(36,37,56)]
DF.t <- DF.t[c(14,20,23,26,29,32,35,38,41,44,47,50,53)]
DF.t = scale(DF.t, center = TRUE, scale = TRUE)
DF.t <- as.data.frame(DF.t)
reg <- lm(BMIDifference ~., DF.t)
```

MODEL INFO:
Observations: 7138
Dependent Variable: BMIDifference
*Type:* OLS linear regression

MODEL FIT:
$F(12,7125) = 59.99$, $p = 0.00$
$R^2 = 0.09$
Adj. $R^2 = 0.09$

Standard errors: OLS

|	Est.	2.5%	97.5%	t val.	p	partial.r	part.r
(Intercept)	0.00	-0.02	0.02	0.00	1.00		
Sum52mealslogged	-0.05	-0.11	0.01	-1.71	0.09	-0.02	-0.02
Sum52greenfoods	-0.00	-0.04	0.03	-0.31	0.76	-0.00	-0.00
Sum52exerciselogged	-0.05	-0.09	-0.01	-2.52	0.01	-0.03	-0.03
Sum52.minutesexercise	-0.00	-0.04	0.04	-0.04	0.97	-0.00	-0.00
Sum52stepsrecorded	0.06	0.02	0.09	3.49	0.00	0.04	0.04
Sum52.weighins	0.25	0.22	0.27	17.66	0.00	0.20	0.20
Sum52.articlesassigned	0.14	0.08	0.19	4.65	0.00	0.06	0.05
Sum52articlesread	-0.00	-0.05	0.05	-0.12	0.91	-0.00	-0.00
Sum52grouppost	-0.01	-0.04	0.02	-0.73	0.46	-0.01	-0.01
Sum52groupcomment	0.02	-0.02	0.06	1.04	0.30	0.01	0.01
Sum52grouplikes	0.01	-0.02	0.05	0.74	0.46	0.01	0.01
Sum52messagestocoach	-0.04	-0.07	-0.01	-2.36	0.02	-0.03	-0.03
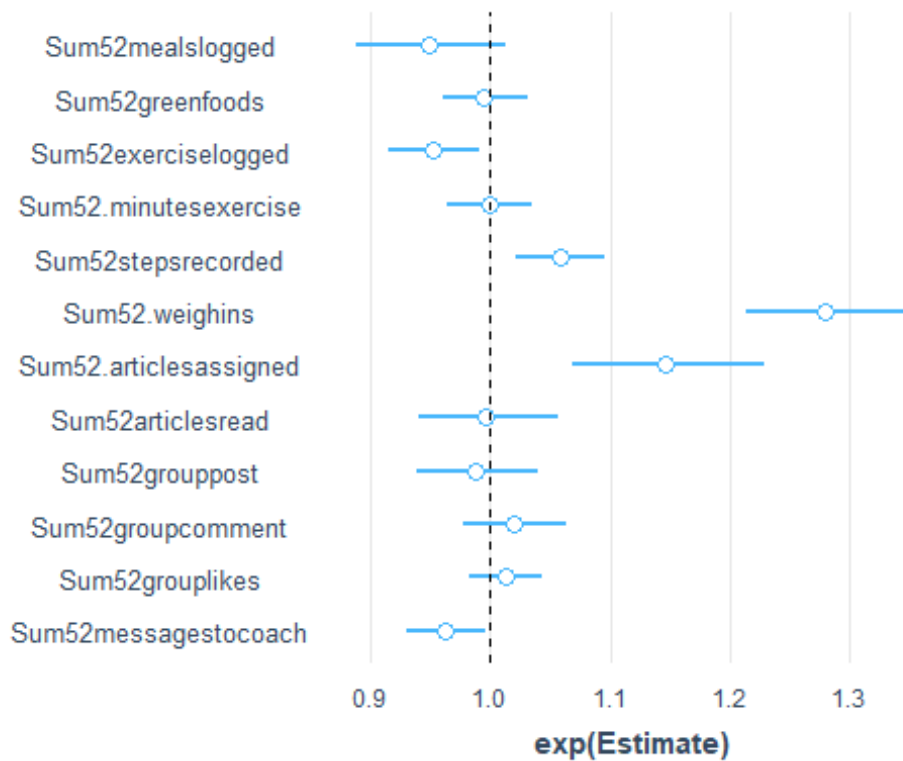
#Regression Tree 5

```
DF <- read.csv("C:/Users/LaoTz/Desktop/DF Articles/WeightLoss.csv", header =
TRUE)
DF <- na.omit(DF)
DF.t <- DF[-c(36,37,56)]
DF.t <- DF.t[c(14,20,23,26,29,32,35,38,41,44,47,50,53)]
model <- train(
  BMIDifference ~., DF.t, method = "ctree",
  trControl = trainControl("cv", number = 10),
  tuneGrid = expand.grid(mincriterion = 0.95)
)
model$results

##   mincriterion     RMSE  Rsquared      MAE    RMSESD RsquaredSD      MAESD
## 1         0.95 2.929279 0.1028164 1.519079 0.2554752 0.03035297 0.05496009
```

Tree Model
```
plot(model$finalModel, type = "simple")
```
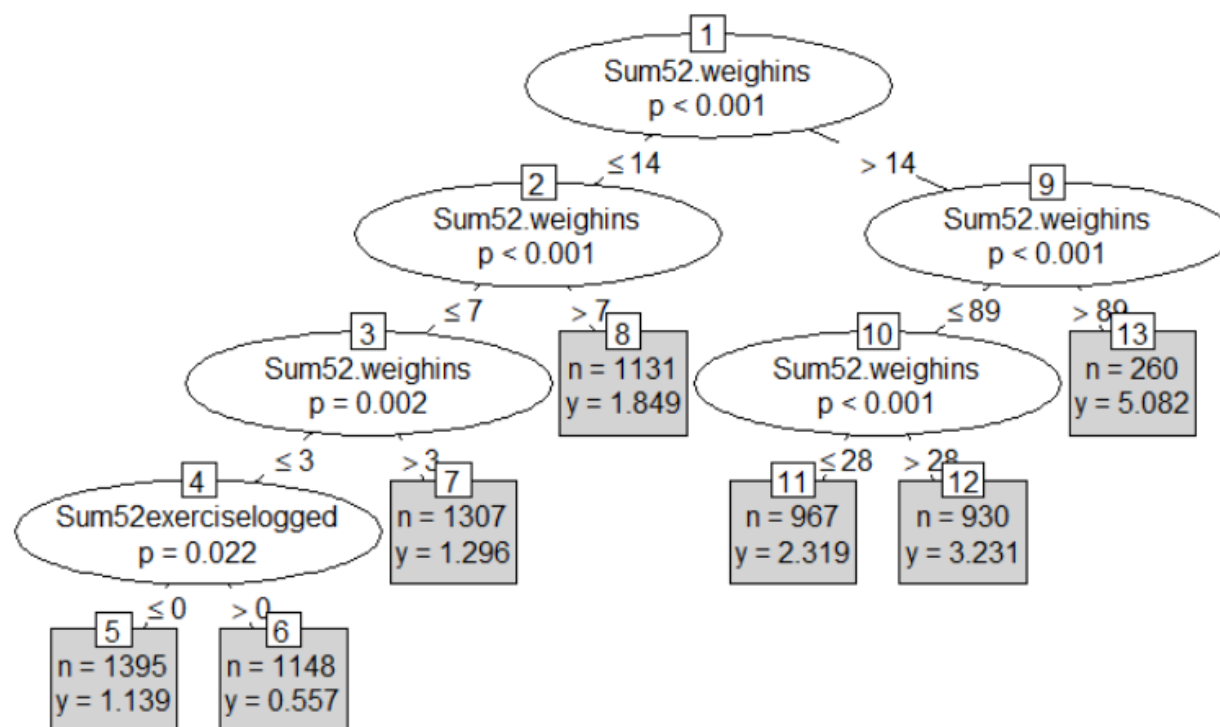
User Engagement and Change in BMI for 7 Terminal Nodes (Left to Right)

1	Diff BMI avg 1.14 = **WK 52 Weigh-ins ≤ 3, WK 52 Exercise Logged ≤ 0**
2 Low	Diff BMI avg 0.56 = **WK 52 Weigh-ins ≤ 3, WK 52 Exercise Logged > 0**
3	Diff BMI avg 1.30 = **WK 52 Weigh-ins > 3 & ≤ 7**
4	Diff BMI avg 1.85 = **WK 52 Weigh-ins > 7 & ≤ 14**
5	Diff BMI avg 2.32 = **WK 52 Weigh-ins > 14 & ≤ 28**
6	Diff BMI avg 3.23 = **WK 52 Weigh-ins > 28 & ≤ 89**
7 High	Diff BMI avg 5.08 = **WK 52 Weigh-ins > 89**

Random Forest 5

```
DF <- read.csv("C:/Users/LaoTz/Desktop/DF Articles/WeightLoss.csv", header =
TRUE)
DF <- na.omit(DF)
DF.t <- DF[-c(36,37,56)]
DF.t <- DF.t[c(14,20,23,26,29,32,35,38,41,44,47,50,53)]
rf <- randomForest(BMIDifference ~ ., data = DF.t, ntree = 25,
                   mtry = 4, nodesize = 5, importance = TRUE)
```

Signicance Testing

```
rf.perm <- rf.significance(rf, DF.t, q = 0.99, p = 0.05, nperm=99, ntree=25)
rf.perm
```

```
## Number of permutations:  99
## p-value:  0.01
## Model signifiant at p = 0.01
##   Model R-square:  -0.08986705
##   Random R-square:  -0.2722119
##   Random R-square variance:  0.0004472861
```

Variable Importance Plot

```r
varImpPlot(rf, type = 1, main = "BMI Diff")
```

BMI Diff